Comp. Biochem. Physiol. Vol. 103B, No. 3, pp. 505–509, 1992
Printed in Great Britain

0305-0491/92 $5.00 + 0.00

MINI REVIEW

A FAMILY OF RETINAL S-ANTIGENS (ARRESTINS) AND THEIR GENES: COMPARATIVE ANALYSIS OF HUMAN, MOUSE, RAT, BOVINE AND *DROSOPHILA*

TOSHIMICHI SHINOHARA,* TAKANOBU KIKUCHI, MASAHIKO TSUDA and KUNIHIKO YAMAKI
Molecular Biology Section, Laboratory of Retinal Cell and Molecular Biology, National Eye Institute, National Institutes of Health, Bethesda, MD 20892, U.S.A. (Tel. 301 496-7799; Fax 301 402-0750)

(*Received* 2 *January* 1992)

Phototransduction, which converts light energy into neuronal impulse, takes place in the rod outer segments of retinal photoreceptor cells. These cells contain several highly specialized proteins such as rhodopsin, cyclic guanosine 3′,5′-monophosphate (cGMP), phosphodiesterase (PDE), Na^+ channels protein and rhodopsin kinase that mediate phototransduction cascade (Stryer, 1988; Chabre and Deterre, 1989).

One of the major soluble proteins in photoreceptor cells, S-antigen (S-Ag; 48-kDa protein or arrestin) is reported to have an inhibitory role in the activated phototransduction cascade. Although the exact mechanism of the inhibition is not known, it is postulated that S-Ag binds to the photoexcited phosphorylated rhodopsin and quenches the elevated light-dependent cGMP-PDE activity (Kuhn *et al.*, 1984; Kuhn and Wilden, 1987; Bennett and Sitaramayya, 1988). Alternatively, S-Ag may directly inhibit cGMP-PDE activation (Zuckerman and Cheasty, 1986; Fukuda *et al.*, 1990). More recently it has been reported that it is an inhibitor of the rhodopsin phosphatase (Palczewski *et al.*, 1989).

S-Ag is also a highly antigenic protein which induces experimental autoimmune uveitis and pinealitis in susceptible animal species (Wacker *et al.*, 1977; Faure, 1980). These autoimmune inflammations have been studied extensively for many years as an animal model for human uveitis (Faure, 1980).

The pineal gland is an organ closely related to the retina and pinealocytes in most species also bear photoreceptor outer segments with their proteins. Interestingly, the mammalian pineal gland is believed to have lost its photoreceptor function as it evolved from a median eye into a secretory organ. In the course of evolution, some major photoreceptor proteins, such as rhodopsin and α-transducin (Tα) might have been eliminated or reduced to minute amounts (van Veen *et al.*, 1986). In contrast, other retinal proteins such as S-Ag and 33-kDa protein (Abe *et al.*, 1990) are major proteins in the mammalian pinealocytes. The role of S-Ag in the mammalian pineal gland is poorly understood. It is believed to bind to the phosphorylated β-adrenergic receptor to inhibit activation of hormonal stimulation (Benovic *et al.*, 1987).

In last few years, cDNAs (Shinohara *et al.*, 1988) and genes for S-Ag have been isolated from several animal species (Yamaki *et al.*, 1990; Tsuda *et al.*, 1991). In addition, a cDNA for β-arrestin (β-A) from rat brain has been characterized (Lohse *et al.*, 1990). The genes for retinal S-Ag of human (Yamaki *et al.*, 1990), mouse (Tsuda *et al.*, 1991), cow, *Drosophila* arrestin (DA) (Smith *et al.*, 1990; Hyde *et al.*, 1990), and 49-kDa arrestin-homolog (49-kDa AH) (Yamada *et al.*, 1990) were studied. Sequence analysis revealed that mammalian S-Ag genes are highly homologous and have 16 exons interrupted by 15 introns spanning approximately 50 kbp in length (Yamaki *et al.*, 1990; Tsuda *et al.*, 1991). The three longest exons, approximately 200 bp, are the 1st, 5th and 16th. The remainder are relatively small, especially the 13th and 15th exons which have only 24 and 10–11 bp, respectively. In contrast, the introns are significantly larger than the exons and the average size of introns is approximately 3 kbp in length. The genes comprise 97% introns and 3% exons. Each splice junction was in good agreement with the GT/AG rule (Breathnach and Chamborn, 1981) for the nucleotides immediately flanking each exon border. In contrast to the mammalian genes, *Drosophila* 49-kDa AH and DA genes consist of three introns with four exons and two introns with three exons, respectively, spanning approximately 2–3 kbp in length (Smith *et al.*, 1990; Hyde *et al.*, 1990; Yamada *et al.*, 1990).

These data may be evaluated from an evolutionary perspective since the first exon encodes with the 5′-noncoding region of mammalian mRNAs, the coding region is covered by the remaining 15 exons with an average size of approximately 80 bp. In

*To whom correspondence should be addressed.

Abbreviations: bp, base pair; CAT, chloramphenicol acetyl-transferase; cGMP, cyclic guanosine 3′,5′-monophosphate; DA, *Drosophila* arrestin; IRBP, interphotoreceptor retinoid-binding protein; kbp, kilo base pair; 49-kDa AH, 49-kDa arrestin homolog; ORF, open reading frame; PDE, phosphodiesterase; S-Ag, S-antigen; Tα, α subunit of transducin.

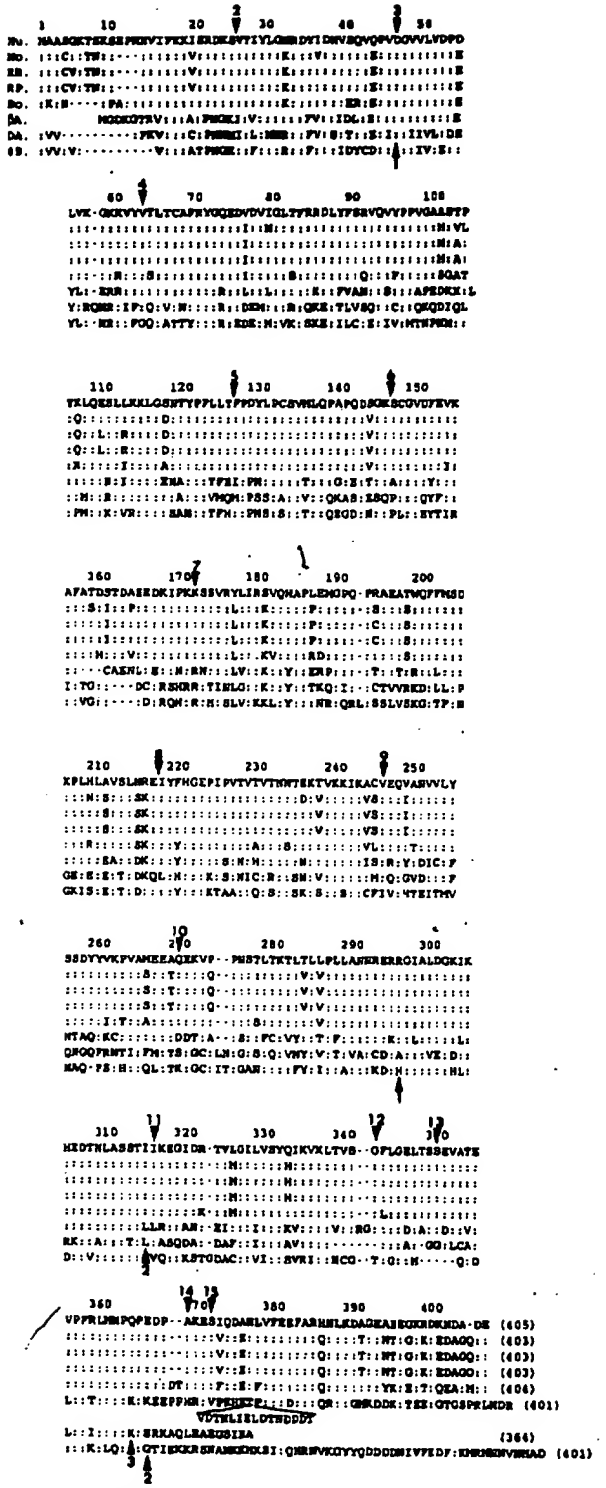

Fig. 1. Comparison of amino acid sequences and splice junctions in a family of S-Ags and their genes. The amino acid sequences from human retina (Hu) (Yamaki *et al.*, 1988), mouse retina (Mo) (Tsuda *et al.*, 1988), rat retina (RR) (Abe *et al.*, 1989), bovine retina (Bo) (Yamaki *et al.*, 1987; Shinohara *et al.*, 1987), rat pineal gland (RP) (Abe and Shinohara, 1990), β-arrestin from rat brain (βA) (Lohse *et al.*, 1990); *Drosophila* arrestin (DA) (Smith *et al.*, 1990; Hyde *et al.*, 1990); and *Drosophila* 49-kDa AH (49) (Yamada *et al.*, 1990). Identical amino acid residues with the human sequence are indicated by colons. An arrowhead indicates splice junctions; arrow pointing upward indicates *Drosophila* and arrow pointing downward indicates human, mouse and part of bovine. The numbers above (mammalian) and below (*Drosophila*) the arrowheads indicate the splice junctions. A hyphen was inserted for creating the maximum alignment.

The number at the end of the sequence indicates the total amino acid residues in each protein.

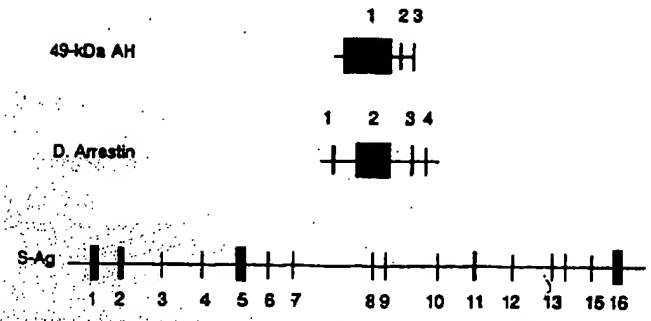

Fig. 2. Gene organization of human, mouse, *Drosophila* S-Ag (DA) and 49-kDa AH. We believe that a common ancestral gene with no intron has evolved into the S-Ag genes by "intron insertion". The black boxes and vertical lines with numbers above (*Drosophila*) and below (mammalian) indicate the exons and horizontal lines indicate introns. Exons 1 and 2 of 49-kDa AH may correspond to exons 1, 2 and 3 of DA, and to exons 2–15 of mammalian S-Ag. Exon 2 and 3 of DA may correspond to exons 3–12 and 13–15 of S-Ag, respectively. The last exons (3 in 49-kDa AH, 4 in DA and 16 in mammalian S-Ag) might diverge.

contrast to the mammalian genes, *Drosophila* DA gene comprises three splice junctions. The first splice junction is identical to mammalian 3rd splice junction. The second junction has only one amino acid residue different from the mammalian 11th junction. The third junction is generally confined to a small region with the mammalian 14th and 15th splice junctions (Fig. 1). However, the 49-kDa AH gene has two splice junctions and the 2nd one contains a *C*-terminus region which is similar to the mammalian 14th and 15th junctions (Fig. 1). A high degree of amino acid sequence conservation in mammal and *Drosophila* was found throughout the S-Ag molecules (50% in DA/S-Ag and 49-kDa AH/S-Ag, 60% in DA/49-kDa AH, and 75% in S-Ag/β-A). The only exception was found at the *C*-terminus regions (Fig. 1). These data suggest that the S-Ag gene may have evolved from a common ancestral intron-poor gene to the genes detected in *Drosophila* with two to three introns (17–19) or mammalian genes with 15 introns by intron insertion (Fig. 2). However, it is possible that S-Ag genes might have evolved from the ancestral intron-rich gene to the intronless gene by intron deletion (Fig. 2). It is of interest to find the gene structure of bacterial S-Ag, such gene structure altered by intron insertion or deletion might play a role in the evolution of S-Ag genes.

The length of the 5′-noncoding sequence in S-Ag mRNA is relatively long (nearly 400 nt). The noncoding sequences on most vertebrate mRNAs fall in the range of 20–100 nt (Kozak, 1978). The effects of noncoding length on translational efficiency are not well understood. Interestingly, two to three small open reading frames (ORF) were found in the 5′-noncoding region of mammalian genes. This feature is not so uncommon; approximately 9% of the vertebrate mRNAs, including oncogenes, have additional upstream ATG codons, and they typically have only one ORF (Mulleller and Hinnebusch, 1986). None of the sequences surrounding the ATG-triplets of these ORF matches Kozak's consensus sequence for an efficient initiator of translation (Kozak, 1978). However, the exact functional role of this configuration and whether or not it produces a protein or if it forms a secondary structure within the small open reading frame remains unknown. Evidence suggests that the level of translation of some mRNAs like the GCN4 in the yeast (Mulleller and Hinnebusch, 1986) and β-adrenergic receptor (Kobilka *et al.*, 1987) are inhibited in part by similar ORF. In the *Drosophila* S-Ag, there is no additional ORF in the 5′-noncoding region of mRNAs.

The 5′-flanking region of the mammalian S-Ag gene, approximately 1.0 kbp in length, had no typical regulatory elements for transcription such as the TATA, CAAT and GC boxes, while a *Drosophila* gene (DA) has TATA and CAAT boxes. We found a poly (A) attaching signal sequence (AATAAA) at the end of the 3′-noncoding sequence in human and mouse S-Ag genes.

Although the 5′-flanking regions of S-Ag genes lack the typical regulatory elements, the 5′-flanking region of the mammalian genes had promoter activity in an *in vitro* transcription assay using a nuclear extract of rat brain. A major transcription start site was found approximately 400 bp upstream from the translation start codon ATG in these three mammalian genes. To study the elements regulating the expression of S-Ag, we generated several transgenic mice expressing the chloramphenicol acetyl transferase (CAT) bacterial reporter gene under the control of a 1,300 bp 5′-flanking segment of the mouse S-Ag gene. Analysis of those mice revealed that CAT activity was present in the retina, pineal gland, lens and brain (Breitman *et al.*, 1991). Immunoblotting, polymerase chain reaction-mediated (PCR) detection of RNA of S-Ag in these tissues, and immunocytostaining of transgenic mouse tissues with antibodies to CAT and S-Ag established that the profile of expression of the transgene corresponded to that of S-Ag; both proteins were detectable in retinal photoreceptor cell, pinealocytes, lens fiber and epithelial cells, the cerebral cortex and cerebellum. This indicates that the DNA 1,300 bp in the 5′-flanking is sufficient for tissue specific expression of CAT and S-Ag [24]. In contrast to the mouse and bovine 5′-flanking region, there are two Alu repeat sequences which were present in the corresponding region of the human gene (Yamaki *et al.*, 1990). Conservation of the nucleotide sequences in the 5′-flanking region

of the mammalian genes was restricted from 1 to −270 bp. Taken together, these results suggest that the transcriptional regulatory element for expression of S-Ag is located from 1 to −270 bp.

The retina and pineal gland in mammalian species have different functions such as phototransduction and the secretion of hormones. The S-Ag and interphotoreceptor retinoid-binding protein (IRBP) are detected at high level in the retina and pineal gland. Promoter sequence for both IRBP and S-Ag genes lack the typical transcriptional regulatory elements (Breitman *et al.*, 1991; Liou *et al.*, 1990). In comparison, the other genes expressed in the photoreceptor cells such as opsin, PDE and Tα genes which contain TATA and CAAT boxes, are not appreciably accumulated in the pineal gland (van Veen *et al.*, 1986). High level protein accumulation in the pineal gland may be in part due to the structural difference of the promoters.

Although S-Ag was previously believed to be a highly specialized single protein, it is now apparent that S-Ags are a family of multiple similar proteins. β-Arrestin, a homolog in the β-adrenergic receptor system has highly homologous sequence (75% similarity) with S-Ag, including the concensus sequences of the phosphoryl binding sites such as G/A-X-X-X-X-G-K (135–141) and D-X-X-G (73–76) were well conserved between S-Ag and β-arrestin (Fig. 1). There is a high degree of homology between two *Drosophila* DA and 49-kDa AH. Moreover, conservation of secondary structure is also high among DA, 49-kDa AH, the mammalian S-Ag, and β-arrestin (Lohse *et al.*, 1990). However, selected differences are apparent, DA lacks a *C*-terminus domain which is postulated to be a rhodopsin binding site (Shinohara *et al.*, 1988). In addition, the presumptive phosphoryl binding sites are conserved well in a single exon in the mammalian species but are diverged in the *Drosophila*. Thus, each protein in a family of S-Ag may have a similar role in its function in different tissues.

REFERENCES

Abe T. and Shinohara T. (1990) S-Antigen from the rat retina and pineal gland have identical sequences. *Exp. Eye Res.* **51**, 111–112.

Abe T., Yamaki K., Tsuda M., Singh V. K., Suzuki S., McKinnon R., Klein D. C., Donoso L. A. and Shinohara T. (1989) Rat pineal S-antigen: sequence analysis reveals presence of α-transducin homologous sequence. *FEBS Lett.* **247**, 307–311.

Bennett N. and Sitaramayya A. (1988). Inactivation of photoexcited rhodopsin in retinal rods: The roles of rhodopsin kinase and 48-kDa protein (Arrestin). *Biochemistry* **27**, 1710–1715.

Benovic J.L., Kuhn H., Weyand I., Codina J., Caron M. G. and Lefkowitz R. J. (1987) Functional desensitization of the isolated β-adrenergic receptor by the β-adrenergic receptor kinase: Potential role of an analog of the retinal protein arrestin (48-kDa protein). *Proc. natn. Acad. Sci. U.S.A.* **84**, 8879–8882.

Breathnach R. and Chamborn O. (1981). Organization and expression of eucaryotic split genes coding for proteins. *Ann. Rev. Biochem.* **50**, 349–383.

Breitman M. L., Tsuda M., Usukura J., Kikuchi T., Zucconi A., Khoo W. and Shinohara T. (1991) Expression of S-antigen in retina, pineal gland, lens, and brain is directed by 5'-flanking sequences. *J. biol. Chem.* **266**, 15505–15510.

Chabre M. and Deterre P. (1989) Molecular mechanism of visual transduction. *Eur. J. Biochem.* **179**, 255–266.

Collin J.-P., Brisson P., Falcon J. and Voisin P. (1986) Multiple cell types in the pineal: functional aspects. In *Pineal and Retina Relationships.* (Edited by O'Brien P. J. and Klein D. C.) pp. 15–32. Academic Press, New York.

Faure J. P. (1980) Autoimmunity and the retina. *Curr. Top. Eye Res.* **2**, 215–302.

Fukuda Y., Yoshizaya T., Saito T., Ohguro S. and Akino T. (1990) Binding of GTP to transducin is not inhibited by arrestin and phosphorylated rhodopsin. *FEBS Lett.* **261**, 419–422.

Hyde D. R., Mecklenburg K. L., Pollock J. A., Vihtelic T. S. and Benzer S. (1990) Twenty *Drosophila* visual system cDNA clones: one is a homolog of human arrestin. *Proc. natn. Acad. Sci. U.S.A.* **87**, 1008–1012.

Kobilka B. K., MacGregor C., Daniel K., Kobilka T. S., Caron M. G. and Lefkowitz R. J. (1987) Translational activity and regulation of human β2-adrenergic receptors expressed in *Xenopus* Oocytes. *J. biol. Chem.* **262**, 15796–15802.

Kozak M. (1978) How do eucaryotic ribosomes select initiation regions in messenger RNA? *Cell* **15**, 1109–1123.

Kuhn H. and Wilden U. (1987) Deactivation and photoactivated rhodopsin by rhodopsin-kinase and arrestin. *J. Recept. Res.* **7**, 283–298.

Kuhn H., Hall S. W. and Wilden U. (1984). Light-induced binding of 48-kDa protein to photoreceptor membranes is highly enhanced by phosphorylation of rhodopsin. *FEBS Lett.* **176**, 473–478.

Liou G., Geng L. I., Alubaidi M. R., Matragoon S., Hanten G., Baehr W. and Overbeek P. K. (1990). Tissue-specific expression in transgenic mice directed by the 5'-flanking sequences of the human gene encoding interphotoreceptor retinoid-binding protein. *J. biol. Chem.* **265**, 8373–8376.

Lohse M. J., Benovic J. L., Codina M. G. and Lefkowitz R. J. (1990) β-Arrestin: a protein that regulates β-adrenergic receptor function. *Science* **248**, 1547–1550.

Mulleller P. and Hinnebusch A. G. (1986) Multiple upstream AUG codons mediate translational control of GCN4. *Cell* **45**, 201–207.

Palczewski K., McDowell J. H., Jakes S., Ingebritsen T. S. and Hargrave P. A. (1989) Regulation of rhodopsin dephosphorylation by arrestin. *J. biol. Chem.* **264**, 1–4.

Shinohara T., Dietzschold B., Craft C. M., Wistow G., Early J. J., Donoso L. A., Horwitz J. and Tao R. (1987) Primary and secondary structure of bovine retinal S-antigen (48-kDa protein). *Proc. natn. Acad. Sci. U.S.A.* **84**, 6975–6979.

Shinohara T., Donoso L. A., Tsuda M., Yamaki K. and Singh V. K. (1988). S-Antigen: Structure, function and experimental autoimmune uveitis (EAU). In *Progress in Retinal Research* (Edited by Osborne N. and Chader J.) Vol. 8, pp. 51–66. Pergamon Press, Oxford.

Smith D. P., Shieh B.-H. and Zuker C. S. (1990) Isolation and structure of an arrestin gene from *Drosophila*. *Proc. natn. Acad. Sci. U.S.A.* **87**, 1003–1007.

Stryer L. (1988) Cyclic GMP cascade of vision. Cold Spring Harb. *Symp. quant. Biol.* **53**, 283–294.

Tsuda M., Syed M., Bugra K., Whelan J. P., McGinnis J. F. and Shinohara T. (1988) Structure analysis of mouse S-antigen. *Gene* **73**, 11–20.

Tsuda M., Kikuchi T., Yamaki K. and Shinohara T. (1991) The mouse S-antigen gene: comparison with human and *Drosophila*. *Eur. J. Biochem.* **200**, 95–101.

van Veen T., Ostholm T., Gietschik P., Spiegel A., Somers R., Korf H. W. and Klein C. D. (1986) α-Transducin immunoreactivity in retinae and sensory pineal organs of adult vertebrates. *Proc. natn. Acad. Sci. U.S.A.* **83**, 912–916.

Wacker W. B., Donoso L. A., Kalsow C. M., Yankeelov J. A. Jr. and Organisciak D. T. (1977) Experimental allergic uveitis, isolation, characterization, and localization of a soluble uveitopathogenic antigen from bovine retina. *J. Immunol.* **119**, 1949–1958.

Yamada T., Takeuchi Y., Komori N., Kobayashi H., Sakai Y., Hotta Y. and Matsumoto H. (1990) A 49-kilodalton phosphoprotein in the *Drosophila* photoreceptor is an Arrestin Homolog. *Science* **248**, 483–486.

Yamaki K., Takahashi Y., Sakuragi S. and Matsubara M. (1987) Molecular cloning of the S-antigen cDNA from bovine retina. *Biochem. biophys. Res. Commun.* **142**, 904–910.

Yamaki K., Tsuda M. and Shinohara T. (1988) The sequence of human retinal S-antigen reveals similarities with α-transducin. *FEBS Lett.* **243**, 39–43.

Yamaki K., Tsuda M., Kikuchi T., Chan K.-H., Huang K.-P. and Shinohara T. (1990) Structural organization of the human S-antigen gene. *J. biol. Chem.* **265**, 20757–20762.

Zuckerman R. and Cheasty J. E. (1986) A 48-kDa protein arrests cGMP phosphodiesterase activation in retinal rod disk membranes. *FEBS Lett.* **207**, 35–41.